

Optimal stability estimates for continuity equations

Christian Seis*

August 23, 2016

Abstract

This review paper is concerned with the stability analysis of the continuity equation in the DiPerna–Lions setting in which the advecting velocity field is Sobolev regular. Quantitative estimates for the equation were derived only recently [38], but optimality was not discussed. In this paper, we revisit the results from [38], compare the new estimates with previously known estimates for Lagrangian flows, e.g. [11], and finally demonstrate how those can be applied to produce optimal bounds in applications from physics, engineering or numerics.

1 Introduction

The linear continuity equation is one of the most elementary partial differential equations. It describes the conservative transport of a quantity by a vector field. We will study this equation in a bounded (Lipschitz) domain¹ Ω in \mathbf{R}^d and denote by $\rho(t, x) \in \mathbf{R}$ and $u(t, x) \in \mathbf{R}^d$ the quantity and the vector field, respectively. For a given initial configuration $\bar{\rho}(x) \in \mathbf{R}$, the Cauchy problem for the continuity equation reads

$$\begin{cases} \partial_t \rho + \nabla \cdot (u\rho) = 0 & \text{in } (0, \infty) \times \Omega, \\ \rho(0, \cdot) = \bar{\rho} & \text{in } \Omega. \end{cases} \quad (1)$$

If the vector field is tangential at the boundary of Ω , which we assume from here on, the quantity ρ is (formally) conserved by the flow:

$$\forall t > 0 : \quad \int_{\Omega} \rho(t, x) \, dx = \int_{\Omega} \bar{\rho}(x) \, dx.$$

Despite its simplicity, the continuity equation plays an important role in fluid dynamics and the theory of conservation laws. In typical applications, ρ represents mass or number density, temperature, energy or phase indicator. In the following, we will frequently refer to ρ as a (possibly negative) mass density, or simply

*Institut für Angewandte Mathematik, Universität Bonn. Email: seis@iam.uni-bonn.de

¹All the results in this work can be extended to the periodic torus or all of \mathbf{R}^d with suitable modifications.

a density. Notice that the vector field has dimensions of length per time, and we will accordingly often refer to u as a velocity field. In fluids applications, u is the velocity of the fluid.

There is a close link between the partial differential equation (PDE) (1) and the ordinary differential equation (ODE)

$$\begin{cases} \partial_t \phi(t, x) = u(t, \phi(t, x)), \\ \phi(0, x) = x. \end{cases} \quad (2)$$

While the PDE represents the Eulerian specification of the flow field, i.e., the description of the dynamics at a fixed location and time, the ODE is the Lagrangian specification, which traces single particles through space and time. The two specifications are in fact equivalent: In the smooth setting, the solution to the continuity equation takes on the nice form

$$\rho(t, \phi(t, x)) \det \nabla \phi(t, x) = \bar{\rho}(x), \quad \text{or simply} \quad \rho(t, \cdot) = (\phi(t, \cdot))_{\#} \bar{\rho}, \quad (3)$$

that is, ρ is the push-forward of $\bar{\rho}$ by the flow ϕ , and a similar formula holds true in the non-smooth setting — as long as (1) and (2) are well-posed. This superposition principle is reviewed in [3, Section 3]. Notice that for any fixed time t , the mapping $\phi(t, \cdot)$ is a diffeomorphism on Ω , whose existence is obtained by the classical Picard–Lindelöf theorem, and $\det \nabla \phi(t, \cdot)$ is the Jacobian determinant, that will be denoted by $J\phi(t, \cdot)$ in the following. The solution ϕ of (2) is called the *flow* of the vector field u .

Out of the smooth setting, well-posedness theory for both the PDE (1) and the ODE (2) is more challenging. We focus on the continuity equation from here on and we start with a suitable concept of generalized solutions in the case where $\bar{\rho} \in L^q(\Omega)$ for some $q \in [1, \infty]$. We call ρ a *distributional solution* to the continuity equation (1) in the time interval $[0, T]$, if it conserves the integrability class of the initial datum, $\rho \in L^\infty((0, T); L^q(\Omega))$, and satisfies

$$-\int_0^T \int_\Omega (\partial_t \zeta + u \cdot \nabla \zeta) \rho \, dx dt = \int_\Omega \bar{\rho} \zeta(0, \cdot) \, dx$$

for any $\zeta \in C_c^\infty([0, T] \times \Omega)$. This distributional formulation is reasonable if $u \in L^1((0, T); L^p(\Omega))$ with $1/p + 1/q = 1$.

In order to prove existence of distributional solutions, we shall impose a condition on the compressibility of the vector field: If u is weakly compressible,

$$(\nabla \cdot u)^- \in L^1((0, T); L^\infty(\Omega)), \quad (4)$$

existence is easily obtained by approximation with smooth functions. Here we have used the superscript minus sign to denote the negative part of the divergence.

The questions of uniqueness and continuous dependence on the initial data are more delicate and have been first answered positively by DiPerna and Lions in their ground breaking paper [20]. Their theory is based on a new solution concept,

the theory of renormalized solutions. A *renormalized solution* is a distributional solution ρ with the property that for any bounded function $\beta \in C^1(\mathbf{R})$ with bounded derivatives and $\beta(0) = 0$ the composition $\beta(\rho)$ satisfies the continuity equation with source

$$\partial_t \beta(\rho) + \nabla \cdot (u \beta(\rho)) = (\nabla \cdot u) (\beta(\rho) - \rho \beta'(\rho))$$

in the sense of distributions. In fact, under the additional assumption that $u \in L^1((0, T); W^{1,p}(\Omega))$, DiPerna and Lions show that distributional solutions *are* renormalized solutions. (The result has been later extended by Ambrosio to vector fields with bounded variation (BV) regularity [2] and recently by Crippa, Nobili and the author to the case where velocity gradient is given by a singular integral of an L^1 function [12].) The advantage of this solution concept is apparent: By choosing $\beta(s)$ as a suitable approximation of $|s|^q$, we obtain by integration over Ω that

$$\frac{d}{dt} \int_{\Omega} |\rho|^q dx = -(q-1) \int_{\Omega} (\nabla \cdot u) |\rho|^q dx \leq (q-1) \|(\nabla \cdot u)^-\|_{L^\infty(\Omega)} \int_{\Omega} |\rho|^q dx,$$

and thus with the help of the Gronwall lemma,

$$\sup_{(0,T)} \|\rho\|_{L^q(\Omega)} \leq \exp \left(\int_0^T \|(\nabla \cdot u)^-\|_{L^\infty(\Omega)} dt \right)^{1-\frac{1}{q}} \|\bar{\rho}\|_{L^q(\Omega)}. \quad (5)$$

By the linearity of the continuity equation, this estimate implies both uniqueness and continuous dependence on the initial data.

Besides proving well-posedness, DiPerna and Lions study stability under approximations of the vector fields and under diffusive perturbations of the equation. (Notice that this gives two different ways of regularizing the PDE.) While qualitative stability estimates are obtained easily via renormalization, the theory fails to provide quantitative stability estimates that capture the rate of convergence of approximate or perturbative solutions to the original one. Such estimates were recently developed in [38].

The aim of the present paper is to revisit the stability estimates from [38] and to reformulate them in a new and optimal way. We will mainly focus on two aspects: Our first intention is to compare the new results with earlier achievements in the theory of Lagrangian flows [11] (that, in fact, strongly inspired the estimates in [38]). Doing so, we hope to convince the reader that the quantities considered in [38] appear naturally in the context of continuity equations. Secondly, we will present applications of the estimates that allow to compute optimal convergence rates in examples of approximate vector fields, zero-diffusivity limits, fluid mixing, and numerical upwind schemes. The last two examples are taken from the studies [37] and [36]. We include these results in order to demonstrate the strength of the estimates from [38] and to underline the intrinsic connection between the respective works (in particular [37]) and the latter. The first example partially extends recent considerations from [15].

We finally like to conclude this introduction by remarking that, as a by-product of the stability estimates, in [38] a new proof of uniqueness is given for (1). This

new proof does not rely on the theory of renormalized solutions but is solely based on the distributional formulation of the equation. In a way, the theory in [38] is the PDE counterpart of the quantitative theory for Lagrangian flows developed by Crippa and De Lellis in [11]. In fact, some of the key estimates were successively transferred from [11] to [38].

Notation: In the following, we will use the shorter notation L^r for the Lebesgue space $L^r(\Omega)$, and similarly $L^r(L^s)$ for $L^r((0, T); L^s(\Omega))$. Further function spaces like $L^1(W^{1,p})$ are defined analogously.

We will omit the domain of integration in the spatial integrals for notational convenience. For instance, we write $\int \cdot dx$ for $\int_\Omega \cdot dx$.

We use the sloppy notation $a \lesssim b$ if $a \leq Cb$ for some constant C that may only depend on the dimension d , the domain Ω or the Sobolev exponent p . We write $a \lesssim_{r_1, \dots, r_n} b$ if C depends in addition on the quantities r_1, \dots, r_n . Finally, we will sometimes use the notation $a \sim b$ if $a \lesssim b$ and $b \lesssim a$.

2 Stability estimates for Lagrangian flows

To motivate our new perspective on the results from [38], we start with recalling some facts from the theory of particles moving in a weakly compressible fluid. The trajectory of a particle moving with the flow is given by the solution of the ODE (2). In the classical setting, when the advecting velocity field u is smooth or at least Lipschitz continuous in the spatial variable, existence and uniqueness of a solution is provided by the Picard–Lindelöf theorem. The Lipschitz regularity also yields simple estimates on the distance of particle trajectories at any time during the evolution. Indeed, as a consequence of the elementary computation

$$\left| \frac{d}{dt} |\phi(t, x) - \phi(t, y)| \right| \leq |u(t, \phi(t, x)) - u(t, \phi(t, y))| \leq \|\nabla u\|_{L^\infty} |\phi(t, x) - \phi(t, y)|$$

and the Gronwall lemma, we easily derive the estimate

$$\exp \left(- \int_0^t \|\nabla u\|_{L^\infty} dt \right) \leq \frac{|\phi(t, x) - \phi(t, y)|}{|x - y|} \leq \exp \left(\int_0^t \|\nabla u\|_{L^\infty} dt \right). \quad (6)$$

Here we have used Rademacher’s identification of Lipschitz functions with the Sobolev class $W^{1,\infty}$, so that $\|\nabla u\|_{L^\infty}$ is the Lipschitz constant of u . This estimate illustrates the well-known fact that two particles transported by the flow can neither converge nor diverge faster than exponentially in time.

This classical result can be equivalently rewritten as

$$- \int_0^t \|\nabla u\|_{L^\infty} dt \leq \log \left(\frac{|\phi(t, x) - \phi(t, y)|}{|x - y|} \right) \leq \int_0^t \|\nabla u\|_{L^\infty} dt, \quad (7)$$

showing that the velocity gradient controls the logarithmic relative distance of two particles. Here “relative distance” refers to the actual distance of particles relative

to their initial distance. We will see in the following that it is the latter perspective rather than the classical one (6) that allows for a generalization to the case of flows for less regular vector fields and also to the Eulerian setting.

Notice that both estimates (6) and (7) contain some information on the regularity of the flow: The flow itself is spatially Lipschitz, uniformly in time, with Lipschitz constant depending on the gradient of u .

Instead of tracing the distance of two different particles in a fluid, we can similarly study the distance of trajectories corresponding to a particle transported by different vector fields: If ϕ and ϕ_k denote the flows associated via (2) to the vector fields u and u_k , respectively, were u_k may be thought of as a Lipschitz continuous perturbation of u , then a computation similar to the one above yields the estimate

$$\log \left(\frac{|\phi(t, x) - \phi_k(t, x)|}{\delta} + 1 \right) \leq \int_0^t \|\nabla u\|_{L^\infty} dt + \frac{1}{\delta} \int_0^t \|u - u_k\|_{L^\infty} dt, \quad (8)$$

for any $\delta > 0$. Therewith, choosing $\delta = \delta_k(t) = \int_0^t \|u - u_k\|_{L^\infty} dt$, we see that

$$\log \left(\frac{|\phi(t, x) - \phi_k(t, x)|}{\delta_k(t)} + 1 \right) \leq \int_0^t \|\nabla u\|_{L^\infty} dt + 1, \quad (9)$$

so that, as before, the velocity gradient controls the logarithmic relative distance of particles moving with two different flows. Observe that $\delta_k(t)$ scales like a length, and can thus be interpreted as the (maximal) distance between the flow fields. Hence, opposed to the situation in (7), we control the distance of particles relative to the distance of vector fields.

Inequality (9) is an estimate on the rate of convergence of trajectories associated with the vector fields u and u_k , if the approximating vector field u_k converges to u in the sense that $\delta_k(t) \rightarrow 0$. The statement then shows that the particle trajectories approach each other with a rate of at least $\delta_k(t)$.

Notice that (8) also implies uniqueness of (2) when the existence of a solution to the ODE is known. Indeed, if $u = u_k$ is spatially Lipschitz and ϕ and ϕ_k are two solutions to (2), the right-hand side of (8) is bounded independently of δ . Hence, choosing δ arbitrarily small we see that ϕ and ϕ_k must be identical.

Out of the smooth setting, the notion of flows for vector fields has to be appropriately generalized. A common generalization is the notion of regular Lagrangian flows that are well-defined if u is merely Sobolev (or even BV) regular in the spatial variable and weakly compressible [20, 2, 11]. The latter is expressed by the requirement that

$$-\infty < \nabla \cdot u(t, x) \quad \text{for a.e. } (t, x) \in (0, T) \times \Omega,$$

cf. (4), which in turn implies that the Jacobi determinant is bounded below:

$$J\phi(t, x) = \exp \left(\int_0^t \nabla \cdot u(t, \phi(t, x)) dt \right) \geq \exp \left(- \int_0^t \|(\nabla \cdot u)^-\|_{L^\infty} dt \right) =: \Lambda. \quad (10)$$

The weak compressibility condition excludes the possibility of infinitely strong sinks in which particles collide in finite time.

Existence, uniqueness and stability of regular Lagrangian flows have been established by DiPerna and Lions in their seminal paper [20] in the case of vector fields with spatial Sobolev regularity (under the assumption that the divergence is uniformly bounded). This theory has been substantially extended to BV vector fields by Ambrosio [2]. We refer the interested reader to the papers [14, 13, 3] for more details and further references, and remark in addition that a comprehensive analysis of the Jacobian is contained in [10].

Interestingly, DiPerna's and Lions's theory for the ODE (2) is built on a well-posedness theory for the associated transport (cf. (17) below) and continuity equations, that is, on the Eulerian (and thus PDE) perspective on particle dynamics. The drawback of the qualitative theory is that no quantitative estimates can be provided. Stability estimates of the type (9) in the DiPerna–Lions setting were derived later by Crippa and De Lellis [11], which are of the form

$$\int \log \left(\frac{|\phi(t, x) - \phi_k(t, x)|}{\delta_k(t)} + 1 \right) dx \lesssim_\Lambda \int_0^t \|\nabla u\|_{L^p} dt + 1, \quad (11)$$

where now

$$\delta_k(t) = \int_0^t \|u - u_k\|_{L^p} dt.$$

We have decided to work with averaged spatial integrals in all formulas in this paper. With that, $\int_0^t \|\nabla u\|_{L^p} dt$ is dimensionless and $\delta_k(t)$ scales like a length, which we interpret, as before, as the distance between the vector fields u and u_k .

The papers [14, 13, 3] provide reviews of DiPerna's and Lions's theory and of Crippa's and De Lellis's contribution.

Obviously, the above result confirms that the earlier principle remains valid: Also in the Sobolev case does the velocity gradient provide control over the logarithmic relative distance of particle trajectories. Moreover, the rate of convergence of the trajectory ϕ_k to the trajectory ϕ is at least of order $\delta_k(t)$ if the latter is tending to zero.

In this weaker setting, the control of the logarithmic distance ceases to hold uniform in space. Nevertheless, the authors are able to deduce local Lipschitz bounds for the generalized flow. (See also [4] for earlier similar results in this direction.) Moreover, uniqueness can be obtained in a way similar to the one outlined above in the case of Lipschitz vector fields.

The quantitative theory of Crippa and De Lellis fails to cover the full range of vector fields considered earlier by DiPerna and Lions [20] and Ambrosio [2]. Instead, the authors have to restrict the setting to Sobolev regular vector fields $u \in L^1((0, T); W^{1,p}(\Omega))$ with $p > 1$. The reason for this is of technical nature: The authors cleverly exploit standard tools from harmonic analysis (maximal functions) whose strong properties just cease to hold if $p = 1$. Stability estimates in the case $p = 1$ (and also the BV case) are still open. On the positive side, in [25], Jabin

manages to extend estimate (11) modulo to a factor of order $o(|\log \delta|)$ to the $W^{1,1}$ setting. This estimate is still strong enough to yield uniqueness and stability — but without rates. A direct proof of uniqueness in the BV setting, that means, without using the uniqueness of the associated partial differential equations as in [2], was obtained (partially) by Jabin [25] and by Hauray and Le Bris [23]. A further extension to the case where the velocity gradient is given by a singular integral is treated by Bouchut and Crippa [5].

It remains to remark that stability estimates in the case $p = 1$ are closely related to a mixing conjecture by Bressan [8]. Indeed, Crippa and De Lellis derive the $p > 1$ analogue of this conjecture in their paper [11] from an estimate similar to (11). See also Subsection 3.5 (or references [37, 24]) for the corresponding result in the Eulerian setting.

3 Stability estimates for continuity equations

In this section, we will present stability estimates in the Eulerian framework that are similar to the ODE theory in [11]. That is, instead of tracing single particles in a fluid, we will study the evolution of macroscopic density functions. Our first intention here is to work out analogies to the Lagrangian framework. For this purpose, we study the case of approximate vector fields in Subsection 3.2. Like the estimates in (9) and (11), the result will be quite general as no relation between the two advecting velocity fields is assumed. In Subsection 3.3, we study convergence rates for the zero-diffusivity limit. Subsection 3.4 is devoted to the convergence order of the numerical upwind scheme. We conclude this paper with an estimate on mixing rates in Subsection 3.5. We start with the introduction of some notation.

3.1 Kantorovich–Rubinstein distance

In order to transfer the Lagrangian stability estimate (11) to the Eulerian specification we need some preparations. The quantity that will replace Crippa’s and De Lellis’s logarithmic trajectory distance is a Kantorovich–Rubinstein distance with logarithmic cost function taken from the theory of optimal transportation and given by

$$\mathcal{D}_\delta(\rho_1, \rho_2) = \inf_{\pi \in \Pi(\rho_1, \rho_2)} \iint \log \left(\frac{|x - y|}{\delta} + 1 \right) d\pi(x, y). \quad (12)$$

Functionals of this type were initially introduced to model minimal costs for transporting mass from one configuration to the other. For two nonnegative distributions ρ_1 and ρ_2 , the set $\Pi(\rho_1, \rho_2)$ consists of all transport plans π that realize this transport, i.e.,

$$\pi[A \times \Omega] = \int_A \rho_1 dx, \quad \pi[\Omega \times A] = \int_A \rho_2 dx,$$

for any measurable set A . The integrand in (12) is the so-called cost function that determines the price for the transport between two points². We refer the interested reader to Villani's monograph [44] for a comprehensive introduction into this topic.

In order to compare this Kantorovich–Rubinstein distance to the trajectory distance considered by Crippa and De Lellis, we notice that in the case where ρ_1 and ρ_2 can be written as push-forwards of the same configuration, which is, for instance, the case if ρ_1 and ρ_2 are advected by different flow fields ϕ_1 and ϕ_2 while having the same initial configuration $\bar{\rho}$ (cf. (3)), then $d\pi = (\phi_1 \otimes \phi_2)_{\#} \delta_{x=y} \otimes d\bar{\rho}$ defines an admissible transport plan in $\Pi(\rho_1, \rho_2) = \Pi((\phi_1)_{\#} \bar{\rho}, (\phi_2)_{\#} \bar{\rho})$. In particular,

$$\mathcal{D}_{\delta}(\rho_1, \rho_2) \leq \int \log \left(\frac{|\phi_1(x) - \phi_2(x)|}{\delta} + 1 \right) \bar{\rho}(x) dx,$$

which means that the Kantorovich–Rubinstein distance $\mathcal{D}_{\delta}(\rho_1, \rho_2)$ is controlled by a weighted variant of Crippa's and De Lellis's logarithmic trajectory distance.

Let us now review some of the properties which make Kantorovich–Rubinstein distances convenient in the study of stability estimates for continuity equations. In fact, what is of importance in our theory is that the quantity $\mathcal{D}_{\delta}(\rho_1, \rho_2)$ constitutes a mathematical *distance* on the space of configurations of same total mass [44, Theorem 7.3], and it metrizes weak convergence [44, Theorem 7.12]. That is,

$$\mathcal{D}_{\delta}(\rho_k, \rho) \longrightarrow 0 \quad \Longleftrightarrow \quad \rho_k \longrightarrow \rho \text{ weakly.}$$

If there exists a sequence of δ_k 's decaying to zero as $k \rightarrow \infty$ and such that $\mathcal{D}_{\delta_k}(\rho_k, \rho)$ is uniformly bounded, the latter thus yields that ρ_k converges weakly to ρ with a rate not larger than δ_k .

A crucial insight the stability analysis of [38] is based on is the dual formulation brought to us in the Kantorovich–Rubinstein theorem

$$\mathcal{D}_{\delta}(\rho_1, \rho_2) = \sup_{\zeta} \left\{ \int \zeta(\rho_1 - \rho_2) dx : |\zeta(x) - \zeta(y)| \leq \log \left(\frac{|x - y|}{\delta} + 1 \right) \right\},$$

cf. [44, Theorem 1.14]. One of its immediate consequences is that $\mathcal{D}_{\delta}(\rho_1, \rho_2)$ is a transshipment cost which only sees the difference between ρ_1 and ρ_2 (“shared mass stays in place”). We can thus write

$$\mathcal{D}_{\delta}(\rho_1, \rho_2) = \mathcal{D}_{\delta}(\rho_1 - \rho_2) \quad \text{or} \quad \mathcal{D}_{\delta}(\rho) = \mathcal{D}_{\delta}(\rho^+, \rho^-),$$

if ρ^+ and ρ^- denote, respectively, the positive and the negative part of ρ . It follows that Kantorovich–Rubinstein distances can be considered as distances between any two *not necessarily nonnegative* configurations with same average.

In [38], the author computed the rate of change of the Kantorovich–Rubinstein distance under the continuity equation with source:

$$\partial_t \rho + \nabla \cdot (u\rho) = \nabla \cdot \sigma.$$

²Concave cost functions are indeed natural in economics applications as they allow to incorporate the *economy of scale* into the mathematical model.

Extending some of the techniques that were developed earlier in [11, 7, 34, 37] in the Lagrangian setting, he found that

$$\left| \frac{d}{dt} \mathcal{D}_\delta(\rho) \right| \lesssim_{\Lambda, \bar{\rho}} \|\nabla u\|_{L^p} + \frac{1}{\delta} \|\sigma\|_{L^1}, \quad (13)$$

cf. [38, Proposition 1], if ρ has zero mean.

3.2 Approximating the vector field

We now consider the situation from (11). We thus let u and u_k be two vector fields in $L^1(W^{1,p})$ satisfying the compressibility condition (10), and we denote by ρ and ρ_k the corresponding solutions to the continuity equation (1), starting with the same initial datum $\bar{\rho}$ in L^q with $1/p + 1/q = 1$. As an immediate consequence of (13) and (5), we obtain our first result.

Theorem 1 ([38]). *If $\delta_k(t)$ denotes the distance between the vector fields u and u_k given by*

$$\delta_k(t) = \int_0^t \|u - u_k\|_{L^p} dt,$$

then it holds

$$\mathcal{D}_{\delta_k}(\rho, \rho_k) \lesssim_{\Lambda, \bar{\rho}} \int_0^t \|\nabla u\|_{L^p} dt + 1.$$

Notice that there is similarity to the control principle we found earlier in the ODE case: The velocity gradient controls the logarithmic relative distance of two configurations.

With regard to the fact that Kantorovich–Rubinstein distances metrize weak convergence, in the situation where $\delta_k \rightarrow 0$, the statement in the theorem now shows that

$$\rho_k \longrightarrow \rho \text{ weakly with rate not larger than } \delta_k.$$

This estimate is sharp as can be seen by the following example suggested by De Lellis, Gwiazda and Świerczewska-Gwiazda [15].

Example 1. *Consider the oscillating vector field $u_k(x) = \sin(2\pi kx)/2\pi k$ on the interval $\Omega = [0, 1]$. Solving the continuity equation with the initial datum $\bar{\rho} = 1$ yields the oscillating solution*

$$\rho_k(t, x) = \frac{1 + \tan^2(\pi kx)}{e^t + e^{-t} \tan^2(\pi kx)},$$

cf. Figure 1. Because u_k converges strongly to zero as $k \rightarrow \infty$, it is clear that the limiting problem is stationary, i.e., $\rho \equiv 1$.

In view of the oscillatory behavior of ρ_k , the convergence to $\rho \equiv 1$ is merely weakly,

$$\|\rho - \rho_k\|_{L^1(L^1)} \not\rightarrow 0.$$

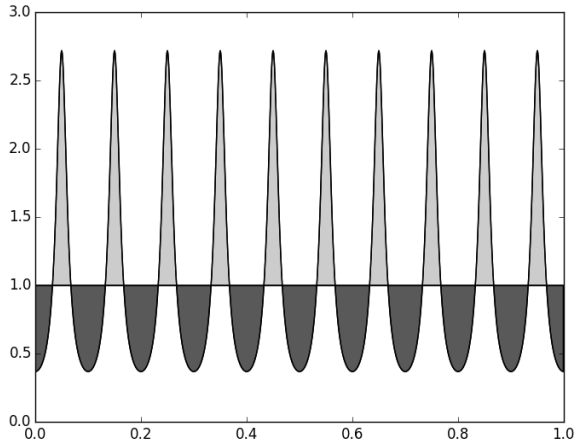


Figure 1: The figure shows the oscillating density ρ_{10} at time $t = 1$. The corresponding Kantorovich–Rubinstein distance $\mathcal{D}_{\delta_{10}}(\rho_{10}, \rho)$ measures the transport between $(\rho_{10} - \rho)^+$ (lightgrey region) and $(\rho_{10} - \rho)^-$ (darkgrey region).

In order to quantify the rate of weak convergence, we notice that

$$\delta_k(t) = \int_0^t \|u - u_k\|_{L^p} dt = \frac{t}{2\pi k} \left(\int_0^1 |\sin(2\pi kx)|^p dx \right)^{1/p} \sim \frac{t}{k}$$

because $u = 0$. By the periodicity and the symmetry of the problem, we furthermore compute

$$\mathcal{D}_{\delta_k}(\rho, \rho_k) = k \mathcal{D}_{\delta_k}(\rho|_{[0, 1/k]}, \rho_k|_{[0, 1/k]}) = \mathcal{D}_{\delta_1}(\rho, \rho_1) \sim_t 1,$$

where we have rescaled length in the last identity.

This example shows that Theorem 1 is sharp in two respects: Firstly, strong convergence of ρ_k to ρ does in general not hold. Secondly, the result captures the correct rate of convergence.

Even though the result in Theorem 1 was already present in [38], the author deduced a weaker stability estimate in order to replace the unwieldy Kantorovich–Rubinstein distance by a standard negative Sobolev norm. In fact, the author proved that

$$\|\rho - \rho_k\|_{W^{-1,1}} \lesssim_{\Lambda, \bar{\rho}, u} \frac{1}{|\log \delta_k(t)|},$$

see [38, Theorem 2]. The new formulation in Theorem 1 has the advantage that it is sharp and naturally extends the analogous estimates in the Lagrangian setting (11).

3.3 The zero-diffusivity limit

In this subsection, we expand the model (1) by a second parallel transport mechanism besides advection: diffusion. Advection-diffusion models are ubiquitous in

thermodynamics, fluid dynamics and engineering, for instance in the context of thermal convection [41], spinodal decomposition [40], or mixing [42]. While convection enhances the efficient transport of particles or fluid parcels over large distances and tends to create sharp gradients in the density (or temperature) distribution, diffusion compensates density (or temperature) differences locally.

We thus consider in the following the Cauchy problem for the advection-diffusion equation

$$\begin{cases} \partial_t \rho_\kappa + \nabla \cdot (u \rho_\kappa) = \kappa \Delta \rho_\kappa & \text{in } (0, \infty) \times \Omega, \\ \rho_\kappa(0, \cdot) = \bar{\rho} & \text{in } \Omega, \end{cases} \quad (14)$$

where κ is the (positive) diffusivity constant. Equipping the equation with the no-flux condition $\nabla \rho_\kappa \cdot \nu = 0$ on the boundary of Ω implies that the evolution is still mass conserving,

$$\int \rho_\kappa(t, x) dx = \int \bar{\rho} dx. \quad (15)$$

We assume furthermore that $\bar{\rho}$ is nonnegative, and so ρ_κ is as a consequence of the maximum principle for (14). Notice that (5) remains valid for ρ_κ , which can be easily seen by testing (14) with ρ_κ^{q-1} .

We are interested in the vanishing diffusivity limit $\kappa \rightarrow 0$. In order to quantify the rate of convergence of solutions of (14) towards solutions of the purely advective model (1), we will make use of a standard decay estimate from the relaxation theory for the diffusion (or heat) equation. A common way to identify the equilibration rate in the diffusive model is by studying the decay behavior of the entropy

$$H(\rho) = \int \rho \log \rho dx.$$

We compute the rate of change of entropy under the evolution (14) using multiple integrations by parts,

$$\begin{aligned} \frac{d}{dt} H(\rho_\kappa) &= \kappa \int \Delta \rho_\kappa \log \rho_\kappa dx - \int \nabla \cdot (u \rho_\kappa) \log \rho_\kappa dx \\ &= - \int \frac{|\nabla \rho_\kappa|^2}{\rho_\kappa} dx - \int (\nabla \cdot u) \rho_\kappa dx, \end{aligned}$$

where in the first equality we have used the fact that the evolution is mass conserving (15). Moreover, since ρ_κ is a nonnegative function, integration in time yields

$$\sup_t H(\rho_\kappa) + \kappa \int_0^t \int \frac{|\nabla \rho_\kappa|^2}{\rho_\kappa} dx dt \leq H(\bar{\rho}) + \left(\int_0^t \|(\nabla \cdot u)^-\|_{L^\infty} dt \right) \|\rho_\kappa\|_{L^\infty(L^1)}.$$

Then, if the initial density has finite entropy, by Hölder's inequality and mass conservation (15),

$$\int_0^t \int |\nabla \rho_\kappa| dx dt \leq \int_0^t \left(\int \rho_\kappa dx \int \frac{|\nabla \rho_\kappa|^2}{\rho_\kappa} dx \right)^{1/2} dt \lesssim_{\bar{\rho}, \Lambda} \sqrt{\frac{t}{\kappa}}.$$

Using the theory developed in [38], cf. (13), and the a priori estimate (5), it can now be shown that

$$\frac{d}{dt} \mathcal{D}_\delta(\rho, \rho_\kappa) \lesssim_{\Lambda, \bar{\rho}} \|\nabla u\|_{L^p} + \frac{\kappa}{\delta} \|\nabla \rho_\kappa\|_{L^1}.$$

Integration in time and a combination of the previous two estimates then yields the following result:

Theorem 2. *Let $\delta_\kappa(t)$ be the diffusion distance per time t , i.e., $\delta_\kappa(t) = \sqrt{t\kappa}$. Then*

$$\mathcal{D}_{\delta_\kappa(t)}(\rho, \rho_\kappa) \lesssim \int_0^t \|\nabla u\|_{L^p} + 1.$$

In other words, the diffusive approximation converges weakly to the unique solution of the continuity equation with a rate not larger than $\sqrt{t\kappa}$. The latter equals approximately the distance a particle can travel by diffusion in time t . Also in this case, a qualitative convergence result was previously established in the work of DiPerna and Lions [20]. To the best of our knowledge, it is for the first time that a convergence rate for the zero-diffusivity limit is obtained.

3.4 Convergence rates for the upwind scheme

The upwind scheme is a numerical scheme for approximating solutions to the continuity equation. The scheme is a finite volume scheme, which means, that the domain is decomposed into control volumes (or cells) of small diameter and the evolving density is approximated by averages over each control volume.

To be more specific, we consider a domain Ω that can be written as a finite union of rectangular boxes. We decompose Ω into a family of rectangular cells with disjoint interiors, $\Omega = \cup_{K \in \mathcal{T}} K$, where \mathcal{T} is the tessellation and K is a translation of the cube $[0, h_1] \times \cdots \times [0, h_d]$. The size h of the tessellation is the maximal edge length, that is,

$$h = \max_{i=1}^d h_i.$$

We suppose that the tessellation is regular in the sense that $h_i \sim h$ for all i . For two neighboring cells $K \sim L$, we denote by $K|L$ the joint boundary. The normal vector on $K|L$ pointing from K to L is denoted by ν_{KL} .

We choose a fixed time step size δt so that the n -th time step reads $t^n = n\delta t$. To guarantee the stability of the explicit scheme, we impose the following Courant–Friedrichs–Lewy (CFL) condition on the time step size:

$$\forall n : \int_{t^n}^{t^{n+1}} \|u\|_{L^\infty} dt \leq h.$$

We are thus assuming in this subsection that $u \in L^1(L^\infty)$.

To approximate the transport term, we consider the net flow from K to L defined by

$$u_{KL}^n = \int_{t^n}^{t^{n+1}} \int_{K|L} u \cdot \nu_{KL} d\mathcal{H}^{d-1} dt.$$

We remark that these quantities are well-defined thanks to the trace estimate for Sobolev functions. Furthermore, the initial configuration of the scheme is the volume average over each $K \in \mathcal{T}$, i.e.,

$$\rho_K^0 = \fint_K \bar{\rho} dx.$$

We are now in the position to define the explicit upwind finite volume scheme for the continuity equation (1),

$$\rho_K^{n+1} = \rho_K^n + \frac{dt}{h} \sum_{L \sim K} (u_{LK}^{n+} \rho_L^n - u_{KL}^{n+} \rho_K^n), \quad (16)$$

where $u_{KL}^{n+} = (u_{KL}^n)^+$. The approximate solution is given by

$$\rho_h(t, x) = \rho_K^n \quad \text{if } (t, x) \in [t^n, t^{n+1}) \times K, \quad K \in \mathcal{T}.$$

See [22, 28] for properties and reference. In the DiPerna–Lions setting, convergence of the scheme, i.e., $\rho_h \rightarrow \rho$ as $h \rightarrow 0$, is proved in [45, 6, 36].

Even though the numerical scheme is formally first order, one observes a break down in the convergence rate to order $1/2$ in the case of non-smooth initial data. In the DiPerna–Lions setting considered here, \sqrt{h} -rates were numerically observed in [6] and [36]. The reason for this lack of convergence is the occurrence of numerical diffusion that smooths out sharp interfaces. Such irregularities, however, are simply transported in the continuous model. In a certain sense, approximate solutions show a behavior similar to those of the advection-diffusion equation (14), where $\kappa \sim h$. It is this similarity that determines the \sqrt{h} -rate of convergence, cf. Theorem 2 above. The effect of numerical diffusion is illustrated in Section 2.4 of [36].

In the case of regular (i.e., at least spatially Lipschitz continuous) vector fields, this break down in the order of convergence is long known. First rigorous results on optimal convergence rates date back to the 1970s, see e.g., [27, 35, 43, 19, 33, 32, 9, 16, 17]. To the best of our knowledge, the only available result in the DiPerna–Lions setting is very recent: In [36], the author establishes jointly with Schlichting an upper bound on the rate of weak convergence that captures the optimal order.

Theorem 3 ([36]). *Let $\delta_h(t)$ be the numerical diffusion distance per time t , i.e., $\delta_h(t) = \sqrt{h \int_0^t \|u\|_{L^\infty} dt}$. Then*

$$\mathcal{D}_{\delta_h}(\rho, \rho_h) \lesssim_{\Lambda, \bar{\rho}} \int_0^t \|\nabla u\|_{L^p} dt + 1.$$

The work [36] builds up not only on the quantitative theory from [38]. A crucial ingredient is a probabilistic interpretation of the upwind scheme suggested by Delarue, Lagoutière and Vauchelet [16, 17]. In fact, in [36], we interpret (16) as a Markov chain, which comes as a time-discretized version of the stochastic differential equation

$$d\psi_t = u(t, \psi_t)dt + \sqrt{2h} dW_t,$$

with a noise term depending on the details of the mesh. In a certain sense, the above equation is the Lagrangian analogue of the advection-diffusion equation (14). It turns out that the noise term determines the \sqrt{h} -rate of convergence.

3.5 Mixing by stirring

In the past years, mixing by stirring attracted much interest in both the applied mathematics and the engineering communities. Mixing refers to the homogenization process of an inhomogeneous substance being stirred by an agent. One of the major goals is the quantification of mixing rates and the design of mixing strategies. In order to optimize mixing strategies, absolute lower bounds on the mixing rate are indispensable. In this subsection, we present a lower bound on mixing by stirring of incompressible viscous fluids that was obtained earlier by the author in [37]. A nice review on the mathematical side of mixing was written by Thiffeault [42].

A natural constraint in the experimental mixing set-up is the amount of mechanical work the engineer is willing to spend in order to overcome viscous friction to maintain stirring. Mathematically, this amounts to limiting the budget of the viscous dissipation rate (or enstrophy) given by $\|\nabla u\|_{L^2}$. In the following, we will slightly generalize this constraint by assuming that $u \in L^1(W^{1,p})$ for some $p > 1$ as in the previous part of this paper.

While our intuition is strong about whether a substance is well mixed or not, the choice of a measure that quantifies the degree of mixedness depends on the mathematical communities. Among fluid dynamists, homogeneous negative Sobolev norms are favored, in particular the $\dot{H}^{-1/2}$ norm [31, 30] and the \dot{H}^{-1} norm [21, 39, 29]. These norms measure oscillations: the larger the length scales, the larger the negative Sobolev norms.

In [37], the author introduces a new mixing measure besides the \dot{H}^{-1} norm: a variant of the Kantorovich–Rubinstein distance introduced earlier in this paper. We accordingly consider

$$M(\rho) = \inf_{\pi \in \Pi(\rho^+, \rho^-)} \exp \left(\iint \log |x - y| d\pi(x, y) \right).$$

Notice that $M(\rho) = \lim_{\delta \rightarrow 0} \exp (\mathcal{D}_\delta(\rho) + (\log \delta) \|\rho\|_{L^1})$. In the case of a two-phase mixture, modeled by $\rho \in \{\pm 1\}$, this distance formally scales as a length, so $M(\rho)$ agrees with the average size of the unmixed regions.

The mixing process can be modeled by the continuity equation (1), which turns into the transport equation

$$\partial_t \rho + u \cdot \nabla \rho = 0 \tag{17}$$

under the assumption that the fluid is incompressible $\nabla \cdot u = 0$, which we shall assume for convenience. For simplicity, we restrict our attention to two-phase mixtures with equal volume fraction, so that

$$|\{x \in \Omega : \rho(t, x) = 1\}| = |\{x \in \Omega : \rho(t, x) = -1\}|$$

for any $t > 0$, or, equivalently, $\int \rho dx = 0$.

In [37], the author derives a lower bound on mixing rates in incompressible viscous fluids, building up on an estimate similar to (13).

Theorem 4 ([37]). *For any $T \geq 0$, it holds that*

$$M(\rho(T, \cdot)) \geq M(\bar{\rho}) \exp \left(-\frac{1}{C} \int_0^T \|\nabla u\|_{L^p} dt \right), \quad (18)$$

where C is a constant depending only on p and d .

This estimate shows the impossibility of perfect mixing, i.e., $\rho \rightarrow 0$ weakly, in finite time. A similar statement has been obtained earlier by Crippa and De Lellis [11] for a certain geometric mixing measure suggested by Bressan [8]. In fact, Bressan conjectures the $p = 1$ analogue of Crippa's and De Lellis's estimate.

It is not difficult to deduce from (18) a lower bound on the decay rate of the \dot{H}^{-1} norm. Indeed, in [37], it is moreover proved that

$$\frac{1}{|\nabla \rho|_{BV}} \lesssim M(\rho) \leq |\nabla^{-1} \rho|_{L^2}, \quad (19)$$

where $|\nabla \rho|_{BV}$ and $|\nabla^{-1} \rho|_{L^2}$ denote, respectively, the homogeneous part of the BV norm, and thus $|\nabla \rho|_{BV} = 2|\partial\{\rho = 1\}|/|\Omega|$, and the homogeneous part of the H^{-1} norm. The first inequality in (19) is an interpolation inequality, whereas the second one follows immediately via Jensen's inequality and the Kantorovich–Rubinstein theorem [44, Theorem 1.14]. Plugging (19) into Theorem 4 yields

$$|(\nabla^{-1} \rho)(T, \cdot)|_{L^2} \gtrsim \frac{1}{|\nabla \rho|_{BV}} \exp \left(-\frac{1}{C} \int_0^T \|\nabla u\|_{L^p} dt \right). \quad (20)$$

A similar decay estimate for the \dot{H}^{-1} norm has been obtained simultaneously by Iyer, Kiselev and Xu [24] by using the geometric results from [11].

Estimates (18) and (20) are sharp. This was proved by Yao and Zlatoš [46] and independently by Alberti, Crippa and Mazzucato [1]. In fact, in both works, the authors construct explicit mixing flows that saturate the lower bounds from [37] and [24]. Numerical evidence for the optimality of this mixing rate was given earlier in [29].

There is a close relation between Theorem 4 and the lower bound in (6). (In fact, also the upper bound

$$M(\rho(T, \cdot)) \leq M(\bar{\rho}) \exp \left(\frac{1}{C} \int_0^T \|\nabla u\|_{L^p} dt \right)$$

is valid.) Estimate (18) can be seen as the Eulerian (and Sobolev) analogue of (6), in the sense that in Theorem 4, we compute the distance between the configuration described by the mixing process and the stationary fully mixed state $\rho = 0$. While (6) shows that trajectories cannot converge faster than exponentially in time, the Eulerian analogue shows that different density configurations cannot converge faster than exponentially in time. This observation also underlines the link between mixing and the question of uniqueness for the partial differential equation (1) (or (17)): A

system is perfectly mixing in finite time precisely if solutions to (1) are in general not unique. Notice that in the case of finite time mixing, upon reversing time, one has nontrivial solutions to (1) with zero initial datum. An explicit construction of such an unmixing solution is due to Depauw [18].

It remains to remark that upper bounds on the rates of unmixing (or coarsening) in viscous fluids were obtained in [7, 34]. The analysis in these papers combines (18) and the lower bound of (19) with the Kohn–Otto upper bound method [26].

Acknowledgement

The author thanks André Schlichting for fruitful discussions and for suggesting the entropy approach in Subsection 3.3.

References

- [1] G. Alberti, G. Crippa, and A. L. Mazzucato. Exponential self-similar mixing and loss of regularity for continuity equations. *C. R. Math. Acad. Sci. Paris*, 352(11):901–906, 2014.
- [2] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [3] L. Ambrosio and G. Crippa. Continuity equations and ODE flows with non-smooth velocity. *Proc. Roy. Soc. Edinburgh Sect. A*, 144(6):1191–1244, 2014.
- [4] L. Ambrosio, M. Lecumberry, and S. Maniglia. Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow. *Rend. Sem. Mat. Univ. Padova*, 114:29–50 (2006), 2005.
- [5] F. Bouchut and G. Crippa. Lagrangian flows for vector fields with gradient given by a singular integral. *J. Hyperbolic Differ. Equ.*, 10(02):235–282, 2013.
- [6] F. Boyer. Analysis of the upwind finite volume method for general initial- and boundary-value transport problems. *IMA J. Numer. Anal.*, 32(4):1404–1439, 2012.
- [7] Y. Brenier, F. Otto, and C. Seis. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM J. Math. Anal.*, 43(1):114–134, 2011.
- [8] A. Bressan. A lemma and a conjecture on the cost of rearrangements. *Rend. Sem. Mat. Univ. Padova*, 110:97–102, 2003.
- [9] B. Cockburn, B. Dong, J. Guzmán, and J. Qian. Optimal convergence of the original DG method on special meshes for variable transport velocity. *SIAM J. Numer. Anal.*, 48(1):133–146, 2010.

- [10] M. Colombo, G. Crippa, and S. Spirito. Renormalized solutions to the continuity equation with an integrable damping term. *Calc. Var. Partial Differential Equations*, 54(2):1831–1845, 2015.
- [11] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.
- [12] G. Crippa, C. Nobili, and C. Seis. Uniqueness for continuity equations with vector field whose gradient is given by a singular integral. In preparation.
- [13] C. De Lellis. ODEs with Sobolev coefficients: the Eulerian and the Lagrangian approach. *Discrete Contin. Dyn. Syst. Ser. S*, 1(3):405–426, 2008.
- [14] C. De Lellis. Ordinary differential equations with rough coefficients and the renormalization theorem of Ambrosio [after Ambrosio, DiPerna, Lions]. *Astérisque*, (317):Exp. No. 972, viii, 175–203, 2008. Séminaire Bourbaki. Vol. 2006/2007.
- [15] C. De Lellis, P. Gwiazda, and A. Świerczewska-Gwiazda. Transport equation with integral terms. Preprint arXiv:1602.03193, 2016.
- [16] F. Delarue and F. Lagoutière. Probabilistic Analysis of the Upwind Scheme for Transport Equations. *Archive for Rational Mechanics and Analysis*, 199(1):229–268, jan 2011.
- [17] F. Delarue, F. Lagoutière, and N. Vauchelet. Convergence order of upwind type schemes for transport equations with discontinuous coefficients. Preprint arXiv:1602.05746, 2016.
- [18] N. Depauw. Non-unicité du transport par un champ de vecteurs presque BV. In *Seminaire: Équations aux Dérivées Partielles, 2002–2003*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XIX, 9. École Polytech., Palaiseau, 2003.
- [19] B. Despres. Lax theorem and finite volume schemes. *Math. Comp.*, 73(247):1203–1234, 2004.
- [20] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.
- [21] C. R. Doering and J.-L. Thiffeault. Multiscale mixing efficiencies for steady sources. *Phys. Rev. E*, 74:025301, Aug 2006.
- [22] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000.
- [23] M. Hauray and C. Le Bris. A new proof of the uniqueness of the flow for ordinary differential equations with BV vector fields. *Ann. Mat. Pura Appl. (4)*, 190(1):91–103, 2011.

- [24] G. Iyer, A. Kiselev, and X. Xu. Lower bounds on the mix norm of passive scalars advected by incompressible enstrophy-constrained flows. *Nonlinearity*, 27(5):973–985, 2014.
- [25] P.-E. Jabin. Differential equations with singular fields. *J. Math. Pures Appl. (9)*, 94(6):597–621, 2010.
- [26] R. V. Kohn and F. Otto. Upper bounds on coarsening rates. *Comm. Math. Phys.*, 229(3):375–395, 2002.
- [27] N. N. Kuznecov. The accuracy of certain approximate methods for the computation of weak solutions of a first order quasilinear equation. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 16(6):1489–1502, 1627, 1976.
- [28] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002.
- [29] Z. Lin, J.-L. Thiffeault, and C. R. Doering. Optimal stirring strategies for passive scalar mixing. *J. Fluid Mech.*, 675:465–476, 2011.
- [30] G. Mathew, I. Mezić, S. Grivopoulos, U. Vaidya, and L. Petzold. Optimal control of mixing in Stokes fluid flows. *J. Fluid Mech.*, 580:261–281, 2007.
- [31] G. Mathew, I. Mezić, and L. Petzold. A multiscale measure for mixing. *Phys. D*, 211(1-2):23–46, 2005.
- [32] B. Merlet. L^∞ - and L^2 -error estimates for a finite volume approximation of linear advection. *SIAM J. Numer. Anal.*, 46(1):124–150, 2007/08.
- [33] B. Merlet and J. Vovelle. Error estimate for finite volume scheme. *Numer. Math.*, 106(1):129–155, 2007.
- [34] F. Otto, C. Seis, and D. Slepčev. Crossover of the coarsening rates in demixing of binary viscous liquids. *Commun. Math. Sci.*, 11(2):441–464, 2013.
- [35] T. E. Peterson. A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.*, 28(1):133–140, 1991.
- [36] A. Schlichting and C. Seis. Convergence rates for upwind schemes with rough coefficients. Preprint arXiv:1606.09156, 2016.
- [37] C. Seis. Maximal mixing by incompressible fluid flows. *Nonlinearity*, 26(12):3279–3289, 2013.
- [38] C. Seis. A quantitative theory for the continuity equation. Preprint arXiv:1602.02931, 2016.
- [39] T. A. Shaw, J.-L. Thiffeault, and C. R. Doering. Stirring up trouble: Multi-scale mixing measures for steady scalar sources. *Physica D: Nonlinear Phenomena*, 231(2):143–164, 2007.

- [40] E. D. Siggia. Late stages of spinodal decomposition in binary mixtures. *Phys. Rev. A*, 20(2):595–605, Aug 1979.
- [41] E. D. Siggia. High Rayleigh Number Convection. *Annual Review of Fluid Mechanics*, 26(1):137–168, 1994.
- [42] J.-L. Thiffeault. Using multiscale norms to quantify mixing and transport. *Nonlinearity*, 25(2):R1–R44, 2012.
- [43] J.-P. Vila and P. Villedieu. Convergence of an explicit finite volume scheme for first order symmetric systems. *Numer. Math.*, 94(3):573–602, 2003.
- [44] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [45] N. J. Walkington. Convergence of the discontinuous Galerkin method for discontinuous solutions. *SIAM J. Numer. Anal.*, 42(5):1801–1817 (electronic), 2005.
- [46] Y. Yao and A. Zlatoš. Mixing and Un-mixing by Incompressible Flows. Preprint arxiv:1407.4163, 2014.